

# Consciousness, Causal Oomph, and Meta-Oomph\*

John A. Barnden  
 School of Computer Science  
 University of Birmingham, UK  
 j.a.barnden@cs.bham.ac.uk

## Abstract

In this paper I argue, on the one hand, in support of a radical claim about the relationship between phenomenal consciousness and causation, and, on the other hand, for the need to focus more attention on an aspect of causation itself. The claims are in the context of an anti-Humean and dynamicist stance, in which there is an irreducible dynamism or causal “oomph” in the world, and causation is just this dynamism or oomph. This paper’s claims centre on a notion of meta-causation (a variety of what has been called iterated causation or higher-order causation). Meta-causation in the intended sense can also be called meta-dynamism or meta-oomph. Meta-causation is where causings—instances or portions of the world’s overall causation/dynamism/oomph—are themselves causes or partial causes. That is, the very *causing* of B by A can *itself, directly* be part of what causes something C.

The distinctive claim about phenomenal consciousness is that there are good reasons to think that it and its basic “feels” or qualia are constitutively based on meta-causation. This claim is ultimately aimed at helping to solve the hard problem of consciousness. But there’s no immediate reason to think that meta-causation is confined to its role in consciousness—it may have much broader significance, though this paper cannot go into what forms that significance might take.

Meta-causation or meta-oomph in the context of consciousness has already been at least implicitly raised in discussions of people possibly having direct conscious experience of causation, because that experience, if it is itself a causal matter, arguably involves meta-causation. More particularly, meta-causation is raised as an issue by recent discussions within power theory of whether we have direct conscious experience of powers at work. However, there appears to have been little discussion of the potential implication that causation, as dynamism or oomph, includes meta-dynamism or meta-oomph, and furthermore that its significance could go beyond consciousness. And as regards the link to consciousness, this paper’s claim goes much further than the idea that we might be able consciously to experience causation, and makes meta-causation constitutively foundational for consciousness itself.

The arguments for this claim start from an assumption that (occurrent, phenomenal) consciousness has a minimal core consisting of pre-reflective self-awareness (PRSA) of the sort championed by recent neo-phenomenologists. PRSA is *pre-reflective* in not being mediated by concepts, inference, or anything else that would ordinarily be taken to be a form of thought. But it is *reflexive* in that the conscious entity is reacting to itself. I argue that there are good reasons to think that this pre-reflective self-reactiveness in turn involves meta-causation.

\* *This draft manuscript paper accompanies a poster, Barnden (2018).*

## 1. Introduction

In this paper I argue, on the one hand, in support of a radical claim about the relationship between phenomenal consciousness and causation and, on the other hand, for the need to focus more attention on an aspect of causation itself.<sup>1</sup>

The claims centre on a notion of meta-causation. In my use of the term here<sup>2</sup>, meta-causation is where causings (instances of causation) are themselves causes or partial causes. That is, the very *causing* of B by A can *itself, directly* be the cause, or part of the cause, of something C, going beyond any effect A and B themselves might have. So for instance we're not talking about chained causation, with simply A causing B and then B causing C. Meta-causation as I will describe it is a metaphysically strong, dynamicist version of what has also variously been called higher-order causation<sup>3</sup> and iterated causation.<sup>4</sup>

My distinctive, radical claim about phenomenal consciousness is that there are good reasons to think that it and its basic “feels” or qualia are constitutively based on meta-causation. This claim is ultimately aimed at helping to solve the hard problem of consciousness. Meta-causation itself is thereby also raised into greater metaphysical prominence than it has previously been given. Moreover, there's no principled reason to think that it is confined to its role in consciousness—it may have much broader significance in the world.

As regards causation in general, I adopt an anti-Humean and dynamicist stance,<sup>5</sup> that there's a

---

<sup>1</sup> The arguments go a considerable distance beyond the preliminary ones I have presented in previous work such as Barnden (2014, 2017).

<sup>2</sup> The term meta-causation has also been used with meanings distinct from mine, such as for forms of top-down causation in , e.g., Ellis (2016).

<sup>3</sup> See, e.g., Koons (1998).

<sup>4</sup> See, e.g., Ehring (2009).

<sup>5</sup> See, e.g., Kutach (2014), Maudlin (2007) and Schaffer (2016) for discussion of Humean views and

basic dynamism in the universe that is the metaphysical grounding for the observed regularity of the happenings in the universe across space and time. The regularities aren't just a brute fact. The dynamism may be a matter of the world being governed, in a deep metaphysical sense, by physical laws, or a matter of the exercise of powers as in recent power theories.<sup>6</sup> Also, I assume change and temporal passage are real, and not just a matter of differences between states in a great big spatiotemporal "block" universe.<sup>7</sup> Given these assumptions, I use the term causation as just another name for the dynamism.

Such dynamism or causation has also been called [*causal* or *metaphysical*] *oomph*.<sup>8</sup> A particular instance of causation is then the instance of dynamism or oomph that binds together the states, events, properties or whatever arising throughout the unfolding of some particular spatiotemporal chunk of the world, i.e. arising in some process. I should stress that the process needn't be sharply divisible from the rest of the world, so the causation within it needn't be sharply divisible from the rest of the causation in the world.

In terms of dynamism, meta-causation is where the instance of dynamism in a spatiotemporal region is itself worked upon by dynamism, much as ordinary states are worked upon by dynamism. Dynamism working on dynamism instances can be called *meta-dynamism*; or, if we use the term oomph we get *meta-oomph*.

Let's turn now to phenomenal consciousness. A large assumption I make here is that a basic, minimal core of consciousness is some form of *pre-reflective self-awareness* (PRSA). Here I follow neo-phenomenologists such as Kriegel (2009) and Zahavi (2005), who are inspired by

---

contrary ones.

<sup>6</sup> See especially Maudlin (2007) on laws, and Mumford & Anjum (2011), Groff & Greco (2013) and Jacobs (2017) on power theory.

<sup>7</sup> Amongst existing major metaphysical theories of time, this paper may be most compatible with the Growing Block and Moving Spotlight views (see Forbes 2016 and Deasy 2015 for recent defences), but the matter needs much more investigation.

<sup>8</sup> See, e.g., Demarest (2017), Kutach (2014), Schaffer (2016).

phenomenologists such as Sartre, Husserl, Heidegger and Merleau-Ponty. PRSA involves some non-intellectually-mediated basic form of reflexive reactivity. It's reflexive in involving reactivity of the conscious entity to itself. I'll argue that this self-reactivity in turn involves meta-causation.

The arguments will appeal to the pre-reflectiveness of the reflexivity as well as to the sheer self-reacting. The pre-reflective aspect is that the self-awareness is non-intellectualized and non-conceptual: it isn't a matter of the conscious process *thinking* about itself in any sense of thinking that involves, even unconsciously, what would readily be described as concepts, beliefs, propositions, inferences, etc. Upholders of the centrality of PRSA thus firmly contrast it with conscious introspection in any ordinary sense of that term, where you're consciously thinking about your own thoughts or feelings and thinking of them *as* thoughts or feelings. Indeed, upholders claim that the non-reflective reflexivity in PRSA is the foundation for the very *consciousness* of such reflective introspection.<sup>9</sup> And PRSA would occur in lower, non-introspecting animals if they were conscious in being able to feel pain, for example.

Meta-causation as a part of consciousness has already been at least implicitly raised in discussions of people possibly having direct conscious experience of causation,<sup>10</sup> and more particularly in recent discussions within power theory of whether we have direct conscious experience of powers at work.<sup>11</sup> Arguably, this would mean that the causation within the experiencing is meta-causation, if the directness is taken seriously and the experiencing is itself a causal matter.<sup>12</sup>

---

<sup>9</sup> In particular, upholders repudiate so-called HOT (higher-order thought) theories of consciousness (Carruthers 2011, Rosenthal 1993).

<sup>10</sup> See, e.g., Beebe (2009).

<sup>11</sup> See especially Mumford (2013) and Mumford & Anjum (2011: ch.9), and also Groff (2013).

<sup>12</sup> But meta-causation in my sense should not be confused with [causation arising from] so-called meta-causal powers (Ellis 2013; also called meta-powers by Groff 2013). A meta-causal power creates, confirms, modifies or destroys other power instances, where a power instance is a power as held by a specific entity, or existing at a specific spatiotemporal location. Meta-causal powers, when exercised, do not engage causation instances themselves in causation, but only cause the creation, modification or destruction of what we could call opportunities for causation.

However, there appears to have been little explicit discussion of the potential implication that causation, as dynamism or oomph, includes *meta*-dynamism or *meta*-oomph, and furthermore that its significance could go beyond consciousness. And as regards the link to consciousness, my claim goes much further than the idea that we can consciously experience causation, and makes meta-causation constitutively foundational for consciousness itself.

Despite my concentration on consciousness, I should stress that there's a much more general reason for studying meta-causation, and indeed for taking it to be part of the physical world. If we grant that causation in the sense of dynamism or oomph is metaphysically real, I see no reason not to say that it's *physically* real. The fact that dynamism itself does not explicitly appear in current equations of basic physics needn't detain us: arguably, the fact that this lack has been untroublesome so far may be because meta-dynamism only occurs in special physical contexts that haven't yet been appropriately experimented upon at the basic physical level—notably, within occurrently conscious brains. But, as soon as one allows that causation instances are a physical reality alongside ordinary physical states, events, etc., one should also consider whether they can themselves take part in causation just as those ordinary physical states, etc. can.

Thus, it isn't that consciousness in its weirdness and wonderfulness leads to the weird notion of meta-causation. Rather, a physicalist theory of consciousness should feel free to help itself to a notion that is naturally suggested by taking an anti-Humean physicalist view of the world's regularities.

I'll now present a mix of further assumptions and argumentation supporting the case for meta-causation in consciousness.

## 2. Phenomenal Consciousness and Meta-Causation

### 2.1 Some Assumptions about Consciousness

In this paper, consciousness means the property of being occurrently, phenomenally conscious.

I make an *Objective Physicalist Assumption* that this property is an *objective* aspect of the *physical* world, at least in the sense that (a) it supervenes on the physical world rather than being in a separate realm of being, and (b) it isn't a construction projected onto the world by people.

The next assumption is a *Causal-Processualist* one, that consciousness is at base a property of *processes*, not of creatures, machines, computer programs, or mental states such as of belief or perception. Talk of such entities being conscious is derivative from the process meaning. By a process I continue to mean a genuine process, one with causal binding of its states.<sup>13</sup> It can occupy a small or large interval of time, in principle.

Given the previous assumptions, the "self" of a conscious process in this paper's conception of pre-reflective self-awareness (PRSA) is just that conscious process itself. It needn't be regarded as the whole self of, say, a person containing that process, and it needn't exist for more than a very short time.

So, self-awareness must be based on some sort of reactiveness of the conscious process to itself. And this reactiveness must be causal, because everything in the physical world happens through causation. Now, any process can in a trivial sense be said to be causally reactive to itself, precisely because successive states in the process's trajectory are causally bound to earlier ones. So something richer than that trivial sense must be involved, and will be discussed below.

---

<sup>13</sup> See Dowe (2009) for discussion of [genuine, causal] processes and pseudo-processes.

The next assumption is behind many thought experiments in the consciousness literature, such as where minds are teleported or people are cloned, or neurons in the brain are gradually replaced by copies without affecting consciousness.<sup>14</sup> Now, I'm neutral as to whether a process is the activity of some physical substrate such as a machine or brain, or is instead free-standing in the sense that either the substrate is considered part of the process itself or there's no principled division in the first place between substrates and processes operating upon them. But, in case we do postulate that there are physical substrates separate from the processes running on them, I make an *Assumption of Substrate Insensitivity and Replaceability*. This has two parts:- (i) No state in a conscious process has components that imply the identity of the particular physical matter in the (token) substrate (particular brain, machine, etc.) it's operating upon; and (ii) at any time, any part or the whole of the physical substrate can in principle be replaced by an identical copy without stopping the process being conscious, or changing its qualia, or changing its cognitive states if it has any—i.e., the conscious process itself experiences no change.

In particular, this Assumption prevents a conscious process that is operating on, say, a brain or machine having an essential self-indexicality or *de se* quality with respect to the physical matter in that particular brain or machine. Rather, it can only have such a quality with respect to the *process* itself.

Now we're ready for two arguments supporting meta-causation. (I present a further major argument in a paper under development, but it's too lengthy and complex to include here.) In the following, it may be helpful to bear in mind that I'm taking causation to be something extra in the physical world over and above ordinary physical states. In particular, the causation or oomph that binds together the states of a process's state trajectory does not supervene on that state trajectory. . Also I should mention that while meta-causation in this paper is where causings are causes,

---

<sup>14</sup> See, e.g., Kirk (2005), Bishop (2009).

there's another side of meta-causation where a causing is causally affected instead. This is potentially important in the further development of my ideas, but I'll leave it aside in this paper.

## 2.2 Argument A, using “I’m-a-Genuine-Process” Reflexivity

What is the self-awareness in PRSA, in terms of mechanism? We can get a start on this question by means of the following, arguably reasonable assumption:

*Assumption of Awareness of Own Processuality:* The pre-reflective self-awareness of a conscious process includes reacting to the process's own *genuine processuality* and not (merely) to its own state trajectory.

After all, if being-conscious is a property of *causally-bound* sequences of states and not just sequences of states, one might reasonably hold that the process isn't genuinely aware of itself unless it's aware of itself *as a genuine process*. I wish to stay true to the pre-reflectiveness of PRSA, so I'm not talking here about the process having a *concept* of genuine processuality and applying that concept to itself. Rather, it's just causally reactive to what it is that makes it a genuine process. This is tantamount to the process being reactive to—i.e., causally affected by—its own internal causation or oomph (as well as possibly its own state trajectory).

So you can see that meta-causation is raising its head. However, we could only immediately infer meta-causation is present if we could immediately take the process's causal reactivity to its own causation to be *direct*—i.e., without any mediation at all. If it is direct, then the meta-causation provides a simple, clear-cut meaning to the idea of a process reacting to own processuality, and provides a straightforward mechanism for it. Note here that the effect of the meta-causation on a state of the process could be simple in quality: as far as our present argument is concerned, there's no need, for instance, to postulate anything in the way of a structured



representation of prior causal relations within the process and the states they relate. The sheer presence of the meta-causation constitutes the reacting to own processuality.

However, one might object to the need to assume that the process's causal reactivity to own causation is direct. Perhaps it's enough for the process merely to react *indirectly* to its own causation, via information about its own state trajectory. One conceivable version of this would be for states within the process to keep representations, of some sort, of prior states. But here we seem to be getting into having to have a naturalistic theory of representation that makes it a completely *objective* matter what past states are being represented. There are great difficulties in defining what objective representation in general amounts to, avoiding any element of our own projection of representation relationships onto the world. And indeed, the additional argument in Barnden (under development) is essentially that there are circumstances in which there would be no objective fact of the matter about what prior state a given state-representation was representing and whether the allegedly represented state is part of the particular process under consideration.

But aside from such concerns, the represented states could (and presumably would) be complex, so the representations would presumably need to be highly structured, with any detailed processing of them arguably amounting to complex inference. I don't necessarily mean here a structuring in traditional symbolic terms; it could instead be on the lines of the continuous dynamic-system based self-reacting of trajectories in Fekete & Edelman (2011). But in either case the whole enterprise would arguably violate the spirit of the term "pre-reflective."

There's no knock-down argument here that indirect, and therefore possibly approximate and uncertain reactivity, to own internal causation could not be achieved without violating pre-reflectiveness. However, by comparison, meta-causation provides a much more straightforward way of accounting for reactivity to own processuality.

### 2.3 Argument B, using “I’m-Self-Aware” Reflexivity

The “self” in the self-awareness in a conscious process’s PRSA is that process, and therefore includes that very self-awareness. Since PRSA is meant to be such a key aspect of consciousness, one might reasonably hold that the process isn’t genuinely aware of itself unless it’s aware of its own PRSA. That the process’s self-awareness should constitutively include awareness of its own PRSA is the *Assumption of Awareness of Own Self-Awareness*.

Then, one can reasonably argue that awareness of own self-awareness must include causally reacting to (some of) the process’s own causation that underlies that self-awareness, and doing so in a pre-reflective way. From this point, the considerations are just a more specific version of the considerations in Argument A. They are possibly yet more troublesome, though, because they potentially require the process to discern which part of its own causation underlies self-awareness as opposed to other operations. Yet, the difficulties arise from an arguably more natural and in some sense weaker assumption than the one of awareness of own processuality.

The argument seeks to show that awareness of own self-awareness involves meta-causation that consists of the process being causally affected by the causation underlying its own self-awareness. But the latter causation, by that very argument, itself involves meta-causation. So we have causation by meta-causation, in other words meta-meta-causation. And so on up through more levels of meta, potentially. But we’re not forced into an infinite regress here, because at each level the process needn’t have full awareness of its own self-awareness, so after a few levels awareness could peter out. In principle there’s also the possibility that the regress could stop because of circular meta-causation, but I’ll not pursue this possibility here.

### **3. Final Remark**

There are many comparisons and contrasts that could be drawn between my view of consciousness and well-known views in the literature. By way of contrast I'm thinking especially here of, first, Higher-Order Thought (HOT) theories of consciousness (Carruthers 2011, Rosenthal 1993). I find these implausible as they stand, but they might be improved by incorporating meta-causation. There's also an interesting contrast with Integrated Information Theory (Oizumi, Albantakis & Tononi 2014). Although this is based on patterns of causation it doesn't involve meta-causation in my sense. Some proposals strongly related in some ways to mine are noted in Barnden (2014), although again they don't appear to involve meta-causation.

## References

- Barnden, J.A. (2014). Running into consciousness. *J. Consciousness Studies*, 21(5–6): 33–56.
- Barnden, J.A. (2017). Consciousness is Loopier than You Thought (When You Run into It). Invited Public Lecture at University of Cagliari, Sardinia, May 2017.
- Barnden, J.A. (2018). Consciousness, causal oomph and meta-oomph. Poster presented at the *26th Conference of the European Society for Philosophy and Psychology*, 10–13 Sep. 2018, Rijeka, Croatia.
- Beebe, H. (2009). Causation and observation. In H. Beebe, C. Hitchcock & P. Menzies (Eds), *The Oxford Handbook of Causation*. pp.471–497. Oxford Scholarship Online ( Oxford University Press).
- Bishop, J.M. (2009). A cognitive computation fallacy? cognition, computations and panpsychism, *Cognitive Computation*, 1, pp.221–233.
- Carruthers, P. (2011) Higher-order theories of consciousness, in Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2011 ed.), online at <http://plato.stanford.edu/archives/fall2011/entries/consciousnesshigher/>.
- Deasy, D. (2015). The moving spotlight theory. *Philosophical Studies*, 172, pp.2073–2089.
- Demarest, H. (2017). Powerful properties, powerless laws. In Jacobs (2017), Ch.4.
- Dowe, P. (2009). Causal processes. In *Stanford Encyclopedia of Philosophy* (Spring 2009 ed.), online at <http://plato.stanford.edu/archives/spr2009/entries/causation-process/>.
- Ehring, D. (2009). Causal relata. In H. Beebe, C. Hitchcock & P. Menzies (Eds), *The Oxford Handbook of Causation*. pp.387–413. Oxford: Oxford University Press.
- Ellis, B. (2013). The power of agency. In Groff & Greco (2013), pp.186–206.
- Ellis, G. (2016). *How can physics underlie the mind? Top-down causation in the human context*. Dordrecht: Springer.
- Fekete, T. & Edelman, S. (2011) Towards a computational theory of experience, *Consciousness and Cognition*, 20, pp. 807–827.
- Forbes, G.A. (2016). The growing block's past problems. *Philosophical Studies*, 173, pp.699–709
- Groff, R. (2013). Whose powers? Which agency? In Groff & Greco (2013), pp.207–227.

- Groff, R. & Greco, J. (Eds) (2013). *Powers and capacities in philosophy: the New Aristotelianism*. New York and London: Routledge.
- Jacobs, J.D. (Ed.) (2017). *Causal powers*. Oxford: Oxford University Press.
- Kirk, R. (2005). *Zombies and consciousness*. Oxford: Clarendon Press (Oxford University Press).
- Koons, R.C. (1998). Teleology as higher-order causation: A situation-theoretic account. *Minds and Machines*, 8: 559–585.
- Kriegel, U. (2009). *Subjective consciousness: A self-representational theory*. Oxford: Oxford University Press.
- Kutach, D. (2014). *Causation*. Cambridge, UK: Polity Press.
- Maudlin, T. (2007). *The metaphysics within physics*. Oxford: Oxford University Press.
- Mumford, S. (2013). The power of power. In Groff & Greco (2013), pp.9–24.
- Mumford, S. & Anjum, R.L. (2011). *Getting causes from powers*. Oxford: Oxford University Press.
- Oizumi, M., Albantakis, L. & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology*, 10 (5): e1003588.
- Rosenthal, D.M. (1993). State consciousness and transitive consciousness. *Consciousness and Cognition*, 2 (4), pp.355–363.
- Schaffer, J. (2016). The metaphysics of causation. In E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), online at <http://plato.stanford.edu/archives/fall2016/entries/causation-metaphysics/>.
- Zahavi, D. (2005). *Subjectivity and selfhood: Investigating the first-person perspective*. Cambridge, Mass. and London, UK: MIT Press (a Bradford Book).